ELSEVIER

Journal of Clinical Epidemiology 53 (2000) 875–883

Journal of Clinical Epidemiology

Measuring cell proliferation in the rectal mucosa: comparing bromodeoxyuridine (BrdU) and proliferating cell nuclear antigen (PCNA) assays

Martin Kulldorff[a,*], Lisa M. McShane[a], Arthur Schatzkin[a], Laurence S. Freedman[a], Michael J. Wargovich[b], Cindy Woods[b], Madhu Purewal[b], Randall W. Burt[c], Michael Lawson[d], Donna J. Mateski[e], Elaine Lanza[a], Donald K. Corle[a], Barbara O'Brien[f], James Moler[g]

[a]*National Cancer Institute, Bethesda, MD*
[b]*M D Anderson Cancer Center, Houston, TX*
[c]*University of Utah, Salt Lake City, UT*
[d]*Kaiser Medical Center, Sacramento, CA*
[e]*Walter Reed Army Medical Center, Washington, DC*
[f]*Westat Inc., Rockville, MD*
[g]*Information Management Services, Rockville, MD*

Received 18 March 1999; received in revised form 2 August 1999; accepted 3 September 1999

Abstract

Cell proliferation in the human colorectum can be measured using bromodeoxyuridine (BrdU) or proliferating cell nuclear antigen (PCNA) assays. Using data from the National Cancer Institute's Polyp Prevention Trial, these two assays are compared using correlation coefficients and variance components analysis. Adjusting for fixed as well as for the random effects of between-biopsy and scoring variation, the estimated correlation is 0.46 for the log labeling index and 0.45 for log proliferative height. This is an estimate of the highest correlation that can be achieved by taking multiple biopsies scored by multiple scorers. For single biopsies, the estimated correlation is 0.16 and 0.10, respectively. There are significant differences between the variance components for the two assays. For example, for log labeling index, PCNA has a lower variation between biopsies than BrdU, but higher variation between scorings. When used in a clinical or epidemiological setting, it is important to take multiple biopsies at multiple time points.

Keywords: Colorectal cancer; Colon adenomas; Correlation; Variance components; Biopsies

1. Introduction

Several studies have examined cell proliferation in relation to colon adenomas and colon cancer [1–5], or in relation to consumption of different dietary components [3,6–14]. The basic hypothesis underlying the dietary studies is that specific dietary components may reduce or increase the cell proliferation rate, which in turn may reduce or increase the number of adenomatous polyps, considered to be precursors for most large bowel cancers.

A key aspect of any such study is the method of measurement of cell proliferation. Rectal mucosal biopsies are taken from study participants and those cells that are in the DNA synthesis phase of the cell cycle are labeled and counted. Two popular methods for labeling such cells involve assays using bromodeoxyuridine (BrdU) [15–18], which labels cells in the S-phase, and proliferating cell nuclear antigen (PCNA) [19–21], which labels cells during the late G_1-phase, all of the S-phase, and early G_2-phase [22]. Comparisons between colorectal mucosal cell proliferation measures based on these two assays have been made in animals [23] as well as in humans [19,24–28]. In this article, data from the National Cancer Institute's Polyp Prevention Trial [29] are used to investigate the general agreement between the two assay measurements, and their relative strengths and weaknesses. The investigation deals with the labeling index, the proportion of labeled cells, as well as the proliferative height, which measures the average relative location of labeled cells within a crypt.

The opinions or assertions contained herein are the private views of the authors and are not to be construed as official or as reflecting the views of the Department of the Army or the Department of Defense.

* Corresponding author. Division of Biostatistics, Department of Community Medicine, University of Connecticut School of Medicine, Farmington, CT 06030-6205, USA. Tel.: 860-679-5473; fax: 860-679-5464.

E-mail address: martink@cortex.uchc.edu (Martin Kulldorff)

0895-4356/00/$ – see front matter
PII: S0895-4356(99)00180-8

When comparing cell proliferation measurements based on the BrdU and PCNA assays, we will always compare biopsies taken from the same participant at the same time. There are several reasons why the measurements on the same participant may not be the same:

1. There is inherent biological variation between any two biopsies, whether they are analyzed using the same or different assays. Likewise, there is variation in the scoring of biopsies when two different scorers score the same biopsy, and even when the same biopsy is scored by the same scorer twice.
2. The two assays do not measure exactly the same biological phenomena, as they may label cells in somewhat different stages of the cell cycle. The fact that the proportions of labeled cells are somewhat different for BrdU and PCNA [30] is an indication of this.
3. Separate handling of the BrdU and PCNA assays that affect the biopsy set as a whole, resulting in somewhat different assay concentrations, temperatures, etc., will introduce variability between the two, just as it would if two PCNA or two BrdU biopsy sets were handled separately.

Depending on the purpose, there are different ways of comparing two methods of measurement. There are situations where the actual values are of primary interest, such as when measuring the blood pressure of particular patients. In other situations, only the relative values are important. The latter is the case in most cell proliferation studies, where the interest is in comparing the level of cell proliferation in different groups of individuals classified according to different dietary habits or interventions, or when cell proliferation is compared at different times for the same individual.

In this article, we use correlation to examine the agreement of the measurements based on the two different assays. This is done at three different levels: that of a single scoring of a single biopsy; that of a biopsy set, typically consisting of three different biopsies each scored by two scorers; and that of the true underlying correlation if we could get perfect measurements from each scorer and biopsy. In the last case, the correlation reflects only the variability under points 2 and 3 above, adjusting for the variability under point 1. We consider this to be the biologically most interesting correlation. Both parametric and nonparametric correlations are calculated, and raw unadjusted correlation coefficients are obtained as well as partial correlation coefficients adjusted for fixed and random effects.

We also compare the individual assay-related variance components. McShane *et al*. [30] used the same data to study different variance components when measuring cell proliferation, separately for the two assays. For either assay, a major source of variation was found between visits at different times by the same participant. This visit to visit variability is here dissected into three components, one that is common to the two assays and the other two that depend on the specific assay used. Moreover, the size of the assay-specific visit-to-visit variability may be different for the two assays. The same is true for between participant, between biopsy and between scorer variability. We compare the assay-specific variance components and test whether they are significantly different for BrdU and PCNA.

2. Materials and methods

2.1. Participants and biopsies

Biopsies were taken from 390 different participants in the Intermediate Endpoint Sub-Study of the Polyp Prevention Trial [29]. The Polyp Prevention Trial is an ongoing multicenter randomized dietary intervention study, with the aim to determine whether a low-fat, high-fiber, high-vegetable, and high-fruit eating plan compared with usual diet, reduces the recurrence of large bowel adenomatous polyps. Participants were all 35 years of age or older, with one or more histologically confirmed large bowel adenomatous polyps removed within the previous 6 months. Of the 390 participants, 32% were women. At baseline, the average age was 62 years, 27% had a family history of colon cancer (sibling, child, or parent), 23% were current aspirin users, and 39% took calcium supplements. Ineligibility criteria for the study included invasive carcinoma in any of the polyps removed, surgical removal of polyps, history of familial polyposis or other polyposis syndromes, history of large bowel adenomatous polyp before age 35, history of large bowel cancer, history of histologically confirmed inflammatory bowel disease, history of large bowel resection, weight greater than 150% of recommended level according to the 1983 Metropolitan Life Insurance Tables, ingestion of lipid-lowering drugs in pharmacological doses within the past month, and dietary pattern similar to the intervention plan.

Biopsies were taken at either or both of two time points, when participants entered the trial and at the end of the first year of follow-up. Eight biopsies were taken from each participant at each visit: two were stored for future use, and three each were analyzed using the BrdU assay and the PCNA method. The average number of evaluable biopsies was 2.0 and 2.6, for the two assays, respectively. Each group of biopsies taken from the same individual at the same time is denoted as a biopsy set.

In total, there are 468 different pairs of evaluable BrdU/PCNA biopsy sets from three different clinical centers: the Kaiser Foundation Research Institute in Oakland ($n = 97$), the University of Utah in Salt Lake City ($n = 161$), and the Walter Reed Army Medical Center in Washington DC ($n = 210$). In some cases, the assay could not be evaluated or too few biopsies were taken and then BrdU only or PCNA only was performed. The number of evaluable biopsy sets, by time period, is given in Table 1.

2.2. Cell proliferation assays

For each pair of BrdU/PCNA biopsy sets, all biopsies were taken from the same participant at the same time. The

Table 1
The number of biospy sets at each time point, from a total of 390 distinct participants

Biopsy sets	Baseline	1 year	total
BrdU/PCNA pairs	205	263	468
BrdU only	9	6	15
PCNA only	109	75	184

biopsies were removed from the endoscopy forceps and immediately placed on a strip of bibulous paper and immersed in minimal essential medium (Sigma Chemical Co.). Within 15 min, and with the help of a microscope, the biopsies were oriented on the paper strip so as to maximize exposure to the medium containing BrdU or fixative, respectively.

For the BrdU assays, the paper strips with the biopsies were placed in disposable borosilicate sample vials, with minimal essential medium containing 50 μM BrdU (Sigma Chemical Co.). After that, 2 ml of $95\%O_2/5\%CO_2$ was injected into the tube. The biopsies were then incubated for 1 h at 37°C, with agitation. After the incubation, the medium was gently removed from the sample vial and the vial was then refilled with 70% ethanol. For the PCNA assays, the minimal essential medium was removed after the orientation procedure, and the shipping tube was filled with 70% ethanol.

For both types of assay, the biopsies were batched and put in a central repository, before being shipped to the MD Anderson Cancer Center for analysis and storage. At MD Anderson all biopsies were processed for histological section and embedded in paraffin. Sections of 4 μm thickness were cut from the samples and placed on poly-*l*-lysine-coated slides.

Those sections of a biopsy with well-oriented crypts were immunostained using an anti-BrdU monoclonal antibody (Becton-Dickinson) or an anti-PCNA PC-10 clone (Signet Laboratories Inc.). Exposure to the monoclonal antibodies was assisted by the use of a semiautomated Sequenza device (Scimetric Inc.). The visualization of the labeled cells was achieved by using the immunoperoxidase method with diaminobenzidine as the chromogen.

2.3. The scoring of biopsies

At MD Anderson, each biopsy was scored independently by two or three different scorers, out of a pool of five. For each biopsy, the scorer first determined whether a particular crypt was well-oriented and therefore scorable. This was defined as those crypts for which the base touched the muscularis mucosa, for which a U-shaped pattern of cells could be traced, and for which there was an open lumen at the top of the crypt.

The scorer then counted the number and location of labeled cells within each scorable crypt. This was done by first assigning position zero to an unlabeled cell in the bottom center of the crypt, and then counting a continuous column of cells along each of the two crypt walls, determining which cells were labeled and which were not. The average height of the crypts was 63.6 cells for biopsies/scorings analyzed with BrdU assay and 67.4 cells for those analyzed with PCNA assay. If all three biopsies of a particular assay failed to yield a total of at least eight scorable crypts, new sections were recut from the block. The average number of crypts scored per biopsy was 6.2 for BrdU and 8.7 for PCNA assays.

Two different cell proliferation measures were calculated: the *labeling index* and the *proliferative height of labeled cells*. The former is defined as the total number of labeled cells divided by the total number of cells. This was calculated on two different aggregation levels, that of a single biopsy scored by a single scorer, and that of a biopsy set where each biopsy is scored by more than one scorer.

The relative height of a labeled cell is the position of that cell divided by the height of the crypt. The relative heights of all labeled cells are then averaged for each biopsy as scored by a particular scorer, to obtain the proliferative height for that biopsy and scorer, and over all labeled cells in all scorings of all biopsies, to obtain the proliferative height for a biopsy set.

Analyses were done using the natural logarithms of the labeling index and proliferative height, which were found to be more in accordance with normal distribution assumptions [30].

2.4. Correlation

Correlation can be calculated at different levels. We can compare cell proliferation in biopsies each scored by a single scorer, one measured using BrdU and the other using PCNA. It is also possible to compare the average cell proliferation in two different biopsy sets, taken at the same time from the same participant. The correlation will typically be higher for biopsy sets than for single biopsies, as some of the scorer and biopsy variability is then averaged out. That is, we have better estimates of the true underlying proliferation measures. Taking this process further, it is possible to estimate the correlation of the true proliferation measurements removing all scorer and biopsy variability, using statistical modeling, as described later.

Correlations for biopsy sets have been calculated using both parametric Pearson correlation and nonparametric Spearman rank correlation [31]. The former assumes normally distributed measurement errors, whereas the latter does not make any distributional assumptions. As the Spearman correlation is based solely on the ranks of the data, the result does not depend on whether the measurements are log transformed or not. In addition to the standard unadjusted correlations, partial correlations [32] were calculated, adjusting for visit, clinical center, hour of biopsy (diurnal trend), month of biopsy (seasonal trend), and date of biopsy (linear time trend). Calculations were done using SAS PROC CORR [33].

The reason for calculating partial correlations is to adjust for confounding covariates (fixed effects). For example, suppose that there is no correlation between the two assay

measurements within any of the clinical centers, but that both measurements are generally much higher for one clinical center than the other. For all clinical centers combined, the unadjusted correlation is then high, but the partial correlation, adjusted for clinical center, would be zero, reflecting the zero correlation in each individual group.

A novel and more flexible way to calculate correlation is through a mixed regression model with both fixed and random effects. A mixed model divides the error into different variance components [34]. For our data the components comprised between-participant variation that is common to the two assay types, $\sigma^2_{\text{participant}}$; between-participant variation due to a specific assay, $\sigma^2_{\text{participant:brdu}}$ and $\sigma^2_{\text{participant:pcna}}$ respectively; participant time interaction that is common to both assays, σ^2_{time}; participant-time interaction due to a specific assay, $\sigma^2_{\text{time:brdu}}$ and $\sigma^2_{\text{time:pcna}}$ respectively; between-biopsy variation nested within the same participant and time $\sigma^2_{\text{biopsy:brdu}}$ and $\sigma^2_{\text{biopsy:pcna}}$; and residual error $\sigma^2_{\text{error:brdu}}$ and $\sigma^2_{\text{error:pcna}}$, due to scoring variability within the same biopsy. Because the same biopsy is never analyzed using both BrdU and PCNA assays, there is no common term for between-biopsy variation or residual error.

More fixed effects can be included in the mixed model than what is possible using partial correlations. The fixed effects included in the mixed model are: assay method, treatment (a dietary intervention), visit (baseline vs. 1 year), clinical center, scorer, scoring date, hour of biopsy (diurnal trend), month of biopsy (seasonal trend), treatment by visit interaction, clinical center by assay interaction, visit by scorer by assay interaction, and scorer by scoring date by assay interaction, as well as the two-factor interaction terms nested within the three-way interactions. This is an extention of McShane's model [30], by including BrdU and PCNA assay measurements in the same regression model, with corresponding addition of specific assay-related fixed and random effects. A detailed motivation for the model is found in McShane *et al*. [30]. Calculations were done using SAS PROC MIXED [33] and restricted maximum likelihood.

Based on the estimated variance components from the mixed model, correlation coefficients can be calculated for a wide variety of settings. Let $\sigma^2_{\text{total:brdu}} = \sigma^2_{\text{participant}} + \sigma^2_{\text{participant:brdu}} + \sigma^2_{\text{time}} + \sigma^2_{\text{time:brdu}} + \sigma^2_{\text{biopsy:brdu}} + \sigma^2_{\text{error:brdu}}$ denote the variance of a single biopsy/scoring based on the BrdU assay, incorporating all the random effects in the model, and define $\sigma^2_{\text{total:pcna}}$ similarly. Note that, because fixed effects do not contribute to the variance, they are not part of the equation above, although they are present in the model. Adjusting for those fixed effects, the partial correlation between two biopsies of different assays, each scored by one scorer, is

$$\frac{\sigma^2_{\text{participant}} + \sigma^2_{\text{time}}}{\sqrt{\sigma^2_{\text{total:brdu}}\,\sigma^2_{\text{total:pcna}}}}. \qquad (1)$$

In the denominator, all variance components are present, reflecting the total variability of a biopsy measurement. In the numerator, on the other hand, only those variance components are present that are common for two different biopsies, taken at the same time, and analyzed by two different assays. The larger these variance components in the numerator are, compared to the total variation (reflected in the denominator), the larger is the correlation.

Adjusting again for fixed effects, the partial correlation between two biopsy sets of different assays, each set consisting of three different biopsies, each scored by two different scorers, is

$$\frac{\sigma^2_{\text{participant}} + \sigma^2_{\text{time}}}{\sqrt{\sigma^2_{\text{total:brdu/set}}\,\sigma^2_{\text{total:pcna/set}}}} \qquad (2)$$

where $\sigma^2_{\text{total:brdu/set}} = \sigma^2_{\text{participant}} + \sigma^2_{\text{participant:brdu}} + \sigma^2_{\text{time}} + \sigma^2_{\text{time:brdu}} + \sigma^2_{\text{biopsy:brdu/3}} + \sigma^2_{\text{error:brdu/2}}$ and where $\sigma^2_{\text{total:pcna/set}}$ is defined equivalently. The multiple biopsies and scorers are reflected in the denominators of the two last entries.

The correlation between a biopsy or biopsy set of BrdU and PCNA assays is necessarily reduced by the between-biopsy and between-scorer variability, even though they have nothing to do with the general agreement between the two assays. If this variability increases, the denominator in eqs. one (1) and two (2) increases, thus reducing the correlation. If we had a very large number of biopsies in each biopsy set, each scored by a very large number of scorers, the biopsy and scorer variability would become irrelevant as we could get nearly perfect estimates of the true underlying cell proliferation for a given assay at a given time. By setting the number of biopsies and scorers to infinity, we get an estimate of the pure correlation of the underlying cell proliferation levels for the two assay types. This gives the partial correlation adjusting not only for the fixed effects, but also for the random effects of biopsy (σ^2_{biopsy}) and scorer (σ^2_{error}) variation. This correlation is

$$\frac{\sigma^2_{\text{participant}} + \sigma^2_{\text{time}}}{\sqrt{\sigma^2_{\text{total:brdu}/\infty}\,\sigma^2_{\text{total:pcna}/\infty}}}$$

where $\sigma^2_{\text{total:brdu}/\infty} = \sigma^2_{\text{participant}} + \sigma^2_{\text{participant:brdu}} + \sigma^2_{\text{time}} + \sigma^2_{\text{time:brdu}}$ and where $\sigma^2_{\text{total:pcna}/\infty}$ is defined equivalently.

Test for concordance [35] is a more stringent criteria than correlation for evaluating the agreement between two different measurements, and it should be applied in many situations where correlation is currently used. If the mean indexes for the two assays are very different, that ensures by itself that the concordance is low. For many uses though, the two assays would work equally well as long as one is a perfect linear combination of the other, such as the Celsius and Fahrenheit scales are for temperature. If that is the case is determined by the correlation coefficient, and this is the rationale for using correlation rather than concordance in this study.

2.5. Means

The mean levels of labeling index and proliferative height were calculated on the ordinary rather than on the log scale and by taking the averages over all biopsy scorings.

These numbers are purely descriptive, as a test for difference between assay methods must be done using the mixed model. To test for a significant difference between these means according to assay method, we compared the full mixed model described above to a reduced one, where both the main and interaction terms related to the assay were removed from the model. This is equivalent to assuming that all fixed effects are the same irrespective of assay method. For this comparison, the mixed models were then fitted using maximum likelihood rather than restricted maximum likelihood [36], and the difference between -2 times the maximized log likelihood for the standard model and -2 times the maximized log likelihood for the reduced model was compared with a chi-square distribution with 17 degrees of freedom.

2.6. Dissection of time variability

In addition to fixed time effects, the model contains three different time-related variance components, one reflecting the visit-to-visit variability within the same participant that is common for the two assays, σ^2_{time}, and the other two reflecting the assay-specific visit-to-visit variability specific for the BrdU assay, $\sigma^2_{\text{time:brdu}}$, and PCNA assay, $\sigma^2_{\text{time:pcna}}$, respectively. Estimates of these variance components are compared to those obtained by McShane *et al.* [30], where the temporal variance component includes both of these types of variability but where it is estimated independently for BrdU and PCNA based on two separate models.

2.7. Comparing BrdU and PCNA variance components

To test whether the variance components related to the BrdU assay are significantly different from the equivalent variance components based on the PCNA assay, the residual log likelihoods were compared for two different mixed models. The first model is described in the correlation section, and the assay-specific variance components are allowed to be different for the two assays. For the second model, each pair of different types of variance components were in turn forced to be equal for the two assays, so that, for example, $\sigma^2_{\text{time:brdu}} = \sigma^2_{\text{time:pcna}}$. Negative two times the difference in residual log likelihoods was then compared to a chi-square distribution with 1 degree of freedom, for testing each pair of components.

3. Results

3.1. Means

The mean labeling index and the mean proliferative height vary depending on assay method, clinical center, scorer, hour of biopsy, and month of biopsy [30]. Averaged over the latter four, the mean labeling index was 4.05% for BrdU assays and 4.39% for PCNA assays. The mean proliferative heights were 29.2% and 26.9%, respectively. Comparing BrdU and PCNA assays, these means were significantly different for both log labeling index ($P < 0.0001$) and log proliferative height ($P < 0.0001$).

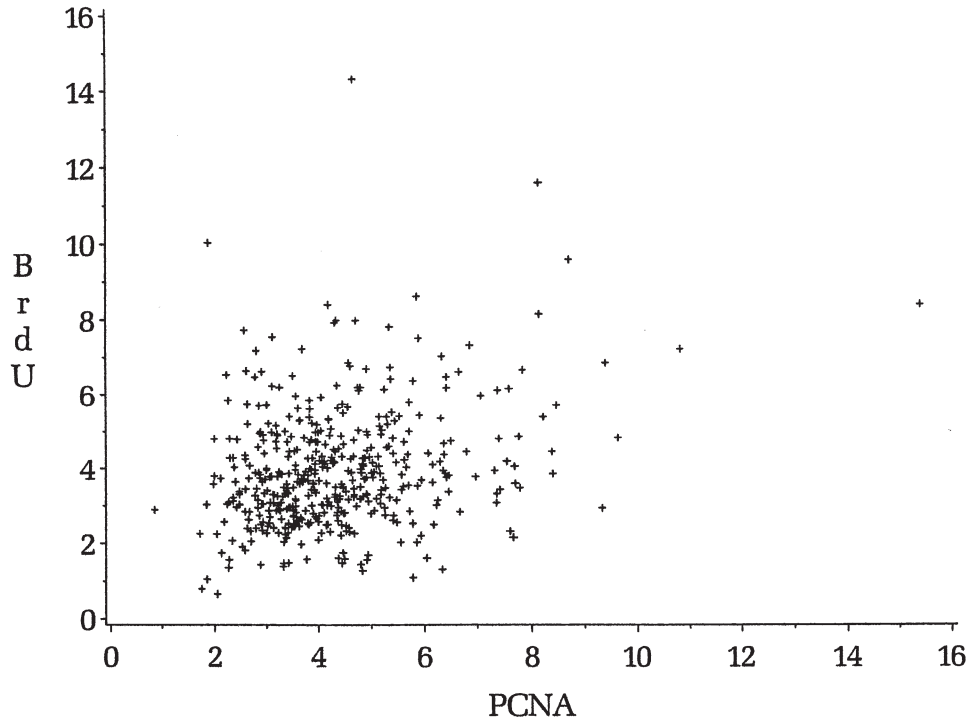

Fig. 1. A scatter plot of the labeling index for biopsy sets using BrdU and PCNA assay respectively.

3.2. Correlation

The BrdU-based labeling index for biopsy sets is plotted against the PCNA-based labeling index in Fig 1, while Fig 2 contains the same plot for proliferative height. The unadjusted Pearson and Spearman rank correlations, based on the 468 matched biopsy sets, are provided in Table 2. The different correlation coefficients are all in the range 0.22–0.28, and in all cases significantly different from zero at significance level 0.0001.

Table 2 also shows the partial Pearson and Spearman rank correlation coefficients. These are adjusted categorically for visit (T_0 or T_1), clinical center, month of biopsy (seasonal trend), and hour of biopsy (diurnal trend), and continuously for date of biopsy (linear yearly trend). The results are very similar to those of the unadjusted correlations, with the coefficients in the 0.22–0.29 range. All are significantly different from zero at the significance level 0.0001.

Correlations based on the mixed model are presented in Table 3. The correlation coefficients for biopsy sets are 0.27 for log labeling index and 0.18 for log proliferative height. These model-based estimates are similar to the Pearson and Spearman correlations presented in Table 2. This is expected. The slight differences are due to estimates of fixed effects that are now based on all observations and not only those for which there is a BrdU/PCNA pair, to additional fixed effects such as scorer that could be adjusted for only in the mixed model, and because the actual biopsy sets do not all contain three biopsies each scored by two scorers as assumed using the mixed model method. Also, the Spearman correlation is a nonparametric correlation coefficient, whereas the mixed model is based on a normal distribution error structure just like the Pearson correlation.

With the mixed model, it is also possible to estimate the correlation between two individual biopsies each scored by one scorer. This correlation is by nature smaller than for biopsy sets, with a coefficient of 0.16 for log labeling index and 0.10 for log proliferative height (Table 3).

From a biological point of view, the most interesting correlation is the one adjusted for the random effects of scorer and biopsy. From Table 3 it can be seen that the estimate of this correlation is 0.46 for log labeling index and 0.45 for log proliferative height.

3.3. Dissection of time variability

McShane *et al*. [30] analyzed the variance components for BrdU and PCNA assays using two separate models. For both assay methods, a major source of variation was found between the two visits at times 1 year apart by the same participant, as shown in the lower part of Table 4.

The time variability could be due either to temporal variation of the cell proliferation as reflected in common by the two assays, or it could be due to temporal variability unique to each assay. The upper part of Table 4 presents the time variation decomposed into these two parts, denoted as “time” and “time:assay,” respectively. It can be seen that the time variability due to assay method can explain about two thirds of the total time variability of the log labeling index and even more for the log proliferative height.

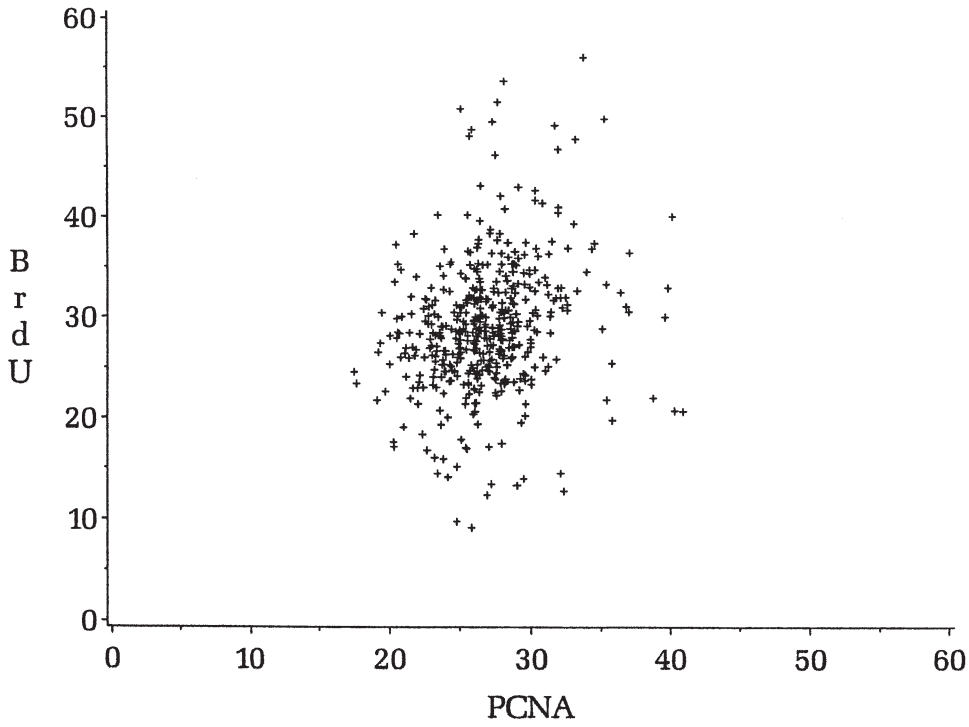

Fig. 2. A scatter plot of the proliferative height for biopsy sets using BrdU and PCNA assay respectively.

Table 2
Pearson and Spearman rank correlations between biospy sets analyzed with BrdU versus PCNA assay

	Log labeling index		Log proliferative height	
	Correlation	P-value	Correlation	P-value
Pearson correlation				
Unadjusted	0.28	<0.0001	0.23	<0.0001
Partial	0.29	<0.0001	0.22	<0.0001
Spearman rank correlation				
Unadjusted	0.22	<0.0001	0.26	<0.0001
Partial	0.23	<0.0001	0.26	<0.0001

The partial correlations are adjusted for visit (baseline vs. 1 year), clinical center, hour of biospy (diurnal trend), month of biopsy (seasonal trend), and date of biospy (yearly trend).

In contrast, for the log labeling index all of the participant variability is common to the two assays, with the assay specific parts estimated to be zero.

3.4. Comparing BrdU and PCNA variance components

The two different assays may differ in their variation over time, in their variability among scorers (error), and in the biopsy variation within the same biopsy set. From Table 5, we can see that the variation from one time to the next is about the same for the BrdU and PCNA assays, and the difference is not statistically significant. In terms of biopsy variation within a biopsy set, the PCNA assay has a much smaller variability than the BrdU based assay, and the difference is statistically significant. In terms of scorer variability, the PCNA assay has smaller variability for the log proliferative height while the BrdU assay has a smaller variability for log labeling index. Both of these are also statistically significant differences.

4. Discussion

4.1. Correlation

Earlier studies comparing BrdU and PCNA assays in the human colorectum used fewer subjects and subject-visits than available from the Polyp Prevention Trial (Table 6). More importantly though, they are based on biopsy sets with different number of crypts and biopsies, and the correlations obtained are then more of a reflection of those numbers than on the agreement/disagreement between assays. The only way to fully get at the latter is to adjust for the between-biopsy and between-scorer variability.

In this study, the correlation between the two assays is very low for single biopsies as well as for biopsy sets, the mixed model estimates being 0.16 and 0.27, respectively, for log labeling index. This is not surprising in light of the larger scorer and biopsy variability (Table 4). The low correlation can be partly mitigated by taking multiple biopsies. If we could take a very large number of biopsies, the correlation would be close to the correlations after adjusting for both fixed and random effects, estimated at 0.46 and 0.45 for log labeling index and log proliferative height, respectively. Hence, these give estimates of the upper bound on the correlation attainable in any study design.

Table 3
Partial correlation coefficients between BrdU and PCNA assays, based on the mixed model variance components analysis

	Log labeling index	Log proliferative height
Adjusted for fixed effects		
Single biopsies	0.16	0.10
Biospy sets	0.27	0.18
Adjusted for fixed and random effects	0.46	0.45

The correlations between BrdU and PCNA assays vary substantially depending on the number of crypts and biopsies obtained for each visit and individual. This is because there is natural variation between crypts, between biopsies, and between scorings. We cannot directly compare correlation estimates from different studies, unless they have been adjusted for these types of random effects, as we did when calculating the 0.46 and 0.45 correlation coefficients. Such adjustments will hopefully be done in future studies, but meanwhile we need to compare the estimates keeping the evaluated number of crypts and biopsies in mind. When known, the average number of evaluated crypts are given in Table 6 together with the correlation estimates from different studies.

The Polyp Prevention Trial took biopsies from participants with at least one adenoma but without cancer, and hence, our results are most directly comparable to the results of Bostick *et al.* [24] and Kilias *et al*. [26]. This article confirms their results: that there is significant correlation. With very few participants, the correlation estimates reported for

Table 4
Variance components analysis with (upper) and without (lower) assay-specific variance components

	Log labeling index		Log proliferative height	
	BrdU	PCNA	BrdU	PCNA
Joint BrdU/PCNA model				
Participant	0.009		0.003	
Participant:assay	0.000	0.000	0.001	0.002
Time	0.022		0.003	
Time:assay	0.041	0.035	0.008	0.004
Biopsy:assay	0.100	0.026	0.035	0.013
Error:assay	0.058	0.087	0.026	0.019
Total	0.230	0.179	0.076	0.044
Separate models [30]				
Participant	0.022	0.006	0.002	0.005
Time	0.046	0.060	0.011	0.006
Biopsy	0.100	0.026	0.035	0.013
Error	0.058	0.087	0.026	0.019
Total	0.226	0.179	0.074	0.043

The upper is based on the combined BrdU/PCNA model presented in this article, while the lower is based on separate models for the two assays [30].

Table 5
Estimated variance components for BrdU and PCNA assays, with 95% confidence intervals, and with a test to see if they are significantly different

	BrdU		PCNA		χ^2 for
	Variance	95% CI	Variance	95% CI	difference
Log labeling index					
Time:assay	0.041	(0.027,0.069)	0.035	(0.024,0.054)	$P = 0.60$
Biopsy:assay	0.100	(0.086,0.118)	0.026	(0.021,0.035)	$P < 0.0001$
Error:assay	0.058	(0.053,0.064)	0.087	(0.081,0.093)	$P < 0.0001$
Log proliferative height					
Time:assay	0.008	(0.003,0.037)	0.004	(0.002,0.018)	$P = 0.33$
Biopsy:assay	0.035	(0.030,0.043)	0.013	(0.011,0.015)	$P < 0.0001$
Error:assay	0.026	(0.024,0.029)	0.019	(0.018,0.021)	$P < 0.0001$

normal subjects by Kubben [19], Weisgerber [28], and Risio *et al*. [27] are rather imprecise, so it is hard to tell whether the difference is due to the type of colon from which the biopsies were taken, or to the imprecision of the estimates.

We found the correlation for log proliferative height to be about the same magnitude as for the log labeling index. Previous studies did not look at correlation for log proliferative height, so no comparisons can be made.

4.2. Dissection of time variability

McShane *et al*. [30] observed that there was considerable time variability within the same participants, for both BrdU and PCNA assays. The conclusion from this article is that most of this variability is related to the particular assay used. Most likely, it is due to some aspect of the assay handling process that is unique to one of the assays, and that is not completely consistent from one visit to the next. It could also be due to some biological or behavioral factor that affects only one of the two assays, such as if a particular eating habit lengthens the time spent in the G-phase of the cell cycle, as this phase is labeled only by the PCNA assay.

4.3. Comparing BrdU and PCNA variance components

The time-related variance specific to the BrdU assay is not significantly different from the time-related variance specific to the PCNA assay, for either log labeling index or log proliferative height. Note that this does not mean that it is a common variance, as the two variances can be the same size without operating in synchronization with each other. The common variance exists in addition to the assay-specific variances, but is smaller than either of those two.

For log labeling index, the estimated error variance is higher for the PCNA assay than for the BrdU assay, while the opposite is true for log proliferative height. The latter may be due partly to the fact that the PCNA assay tends to label more cells per crypt, which would give a more accurate estimate of their average height in the crypt.

The biopsy variability is considerably smaller for the PCNA than the BrdU assay. This may be related to the difference in biopsy handling, as the BrdU technique is more time consuming and complicated, but it is hard to pinpoint the exact cause. The total variance is smaller for the PCNA assay for both labeling index and proliferative height (Table 4). This may indicate a possible advantage of using the PCNA assay.

4.4. Conclusions

From the correlation study it is clear that the cell proliferation measurements based on the two different assays to some extent measure the same phenomena, but the correla-

Table 6
Different estimates of correlations between PCNA and BrdU assays for measuring the log labeling index of cell proliferation in the human colorectum. The correlations are difficult to compare since they depend on the number of crypts and biopsies taken per visit. For the current study, correlation coefficients are given for single biopsies and for biopsy sets, together with the adjusted correlation coefficient. The latter is the most interesting, reflecting the pure correlation between assays after adjusting for the number of crypts and biopsies.

Study	Subjects	Subject-visits	Average crypts/visit		Correlation
			BrdU	PCNA	
Risio et al. [27]	Adenocarcinoma	10	unknown		0.15
Current Study	Adenomas	468	6.2	8.7	0.16
Risio et al. [27]	Multiple Adenomas	6	unknown		0.23
Bostick et al. [24]	Adenomas	133	10–16	10–16	0.24
Current Study	Adenomas	468	14.4	22.6	0.27
Current Study	Adenomas	468	adjusted		0.46
Kilias et al. [26]	Adenomas	20	21	27	0.49
Weisgerber et al. [28]	Normal	17	⩾15	⩾15	0.60
Risio et al. [27]	One Adenoma	14	unknown		0.61
Kubben et al. [19]	Normal	16	⩾6	⩾6	0.63
Risio et al. [27]	Normal	50	unknown		0.70

tion coefficients are low, making it an uncertain marker for either epidemiological studies or clinical use.

Using variance components analysis, Weisgerber *et al*. [28] recommended that at least two biopsies should be taken from each participant to allow a precise individual characterization. Our analysis confirms that recommendation. In light of the large biopsy and scoring variation and the fairly weak assay correlation, we think it is important to take more than two biopsies from each participant at a visit. If future epidemiological studies are conducted using these assays, it is also important to ensure that the number of participants is large. Our general conclusion is that there is a need to obtain more reliable assay methods for measuring cell proliferation.

References

[1] Biasco G, Paganelli GM, Miglioli M, *et al*. Rectal cell proliferation and colon cancer risk in ulcerative colitis. Cancer Res 1990;50:1156–9.

[2] Bleiberg H, Buyse M, Galand P. Cell kinetic indicators of premalignant stages of colorectal cancer. Cancer 1985;56:124–9.

[3] Frommel TO, Mobarhan S, Doria M, *et al*. Effect of beta-carotine supplementation on indices of colonic cell proliferation. J Natl Cancer Inst 1995;6:1781–7.

[4] Risio M, Lipkin M, Candelaresi G, *et al.* Correlation between rectal mucosa cell proliferation and the clinical and pathological features of nonfamilial neoplasia of the large intestine. Cancer Res 1991;51:1917–21.

[5] Terpstra OT, Blankenstein M, Dees J, *et al*. Abnormal pattern of cell proliferation in the entire colonic mucosa of patients with colon adenomas or cancer. Gastroenterology 1987;92:704–8.

[6] Alberts DS, Einspahr J, Ritenbaugh C, *et al*. The effect of wheat bran fiber and calcium supplementation on rectal mucosal proliferation rates in patients with resected adenomatous colorectal polyps. Cancer Epidemiol Biomarkers Prev 1997;6:161–9.

[7] Armitage NC, Rooney PS, Gifford KA, *et al*. The effect of calcium supplements on rectal mucosal proliferation. Br J Cancer 1995;71:186–90.

[8] Baron JA, Tosteson TD, Wargovich MJ, *et al*. Calcium supplementation and rectal mucosal proliferation: a randomized controlled trial. J Natl Cancer Inst 1995;87:1303–7.

[9] Bostick RM, Fosdick L, Wood JR, *et al*. Calcium and colorectal epithelial cell proliferation in sporadic adenoma patients: a randomized, double-blinded, placebo-controlled clinical trial. J Natl Cancer Inst 1995;87:1307–15.

[10] Buset M, Lipkin M, Winawer S, *et al*. Inhibition of human colonic epithelial cell proliferation in vivo and in vitro by calcium. Cancer Res 1986;46:5426–30.

[11] Caderni G, Bianchini F, Russo A, *et al*. Miotic activity in colorectal mucosa of healthy subjects in two Italian areas with different dietary habits. Nutr Cancer 1993;19:263–8.

[12] Einspahr JG, Alberts DS, Gapstur SM, *et al*. Surrogate end-point biomarkers as measures of colon cancer risk and their use in cancer chemoprevention trials. Cancer Epidemiol Biomarkers Prev 1997;6: 37–48.

[13] Lipkin M, Newark H. Effect of added dietary calcium on colonic epithelial-cell proliferation in subjects at high risk for familial colonic cancer. N Engl J Med 1985;313:1381–4.

[14] Stadler J, Stern HS, Sing Yeung K, *et al*. Effect of high fat consumption on cell proliferation activity of colorectal mucosa and on soluble faecal bile acids. Gut 1988;29:1326–31.

[15] Darmon E, Pincu-Hornstein A, Rozen P. A rapid and simple in vitro method for evaluating human colorectal epithelial proliferation. Arch Pathol Lab Med 1990;114:855–7.

[16] Morstyn G, Pyke K, Gardner J, *et al*. Immunohistochemical identification of proliferating cells in organ culture using bromodeoxyuridine and a monoclonal antibody. J Histochem Cytochem 1986;34:697–701.

[17] Potten CS, Kellett M, Roberts SA, *et al*. Measurement of in vivo proliferation in human colorectal mucosa using bromodeoxyuridine. Gut 1992;33:71–8.

[18] Roe DJ, Alberts DS, Wargovich MJ, *et al*. Reproducibility of the measurement of colonic proliferation using bromodeoxyuridine across laboratories. Cancer Epidemiol Biomarkers Prev 1986;5:349–53.

[19] Kubben FJ, Peeters-Haesevoets A, Engels LG, *et al*. Proliferating cell nuclear antigen (PCNA): a new marker to study human colonic cell proliferation. Gut 1994;35:530–5.

[20] Takasaki Y, Deng JS, Tan EM. A nuclear antigen associated with cell proliferation and blast transformation. J Exp Med 1981;154:1899–1909.

[21] Tanaka K, Murata N, Yanai H, Okita K. Immunohistological study on the expression of proliferating cell nuclear antigen (PCNA/cyclin) in human colorectal lesions. Nippon Shokakibyo Gakkai Zasshi 1992;89:493–7.

[22] Kurki P, Ogata K, Tan EM. Monoclonal antibodies to proliferating cell nuclear antigen (PCNA/cyclin) as probes for proliferating cells by immunofluorescence microscopy and flow cytometry. J Immunol Methods 1988;109:49–59.

[23] Richter F, Richter A, Yang K, Lipkin M. Cell proliferation in rat colon measured with bromodeoxyuridine, proliferating cell nuclear antigen, and [H3]thymidine. Cancer Epidemiol Biomarkers Prev 1992;1:561–6.

[24] Bostick RM, Fosdick L, Lillemoe TJ, *et al*. Methodological findings and considerations in measuring colorectal epithelial cell proliferation in humans. Cancer Epidemiol Biomarkers Prev 1997;6:931–42.

[25] Bromley M, Rew D, Becciolini A, *et al*. A comparison of proliferation markers (BrdUrd, Ki-67, PCNA) determined at each cell position in the crypts of normal human colonic mucosa. Eur J Histochem 1996;40:89–100.

[26] Kilias D, Macrae FA, Sharpe K, Young GP. The effect of incubation of rectal biopsies on measures of proliferation using proliferating cell nuclear antigen in comparison with 5-bromo-2-deoxyuridine. Cancer Epidemiol Biomarkers Prev 1997;6:819–24.

[27] Risio M, Candelaresi G, Rossini FP. Bromodeoxyuridine uptake and proliferating cell nuclear antigen expression throughout the colorectal tumor sequence. Cancer Epidemiol Biomarkers Prev 1993;2:363–7.

[28] Weisgerber UM, Boeing H, Nemitz R, *et al*. Proliferation cell nuclear antigen (clone 19A2) correlates with 5-bromo-2-deoxyurdine labeling in human colonic epithelium. Gut 1993;34:1587–92, 1994;35:717–8.

[29] Schatzkin A, Lanza E, Freedman LS, *et al*. The polyp prevention trial: I. Rationale, design, recruitment, and baseline participant characteristics. Cancer Epidemiol Biomarkers Prev 1996;5:375–83.

[30] McShane LM, Kulldorff M, Wargovich MJ, *et al*. An evaluation of rectal mucosal proliferation measure variability sources in the polyp prevention trial: can we detect informative differences among individuals' proliferation measures amid the noise? Cancer Epidemiol Biomarkers Prev 1998;7:605–12.

[31] Spearman C. The proof and measurement of association between two things. Am J Psychol 1904;15:72–101.

[32] Yule GU. On the significance of Bravais' formulae for regression in the case of skew correlation. Proc R Soc Lond B 1897;60:477–89.

[33] SAS v 6.12, SAS Institute Inc., SAS Circle, Box 8000, Cary NC 27512-8000, USA.

[34] Searle SR, Casella G, McCulloch CE. Variance Components. New York: Wiley, 1992.

[35] Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989;45:255–68.

[36] Wolfinger R. A tutorial on the mixed models. In: Cary, NC: Getting Started With PROC MIXED. Cary, NC: SAS Institute Inc., 1992.